Point/Counterpoint

The Reliability of Psychiatric Diagnoses

POINT—OUR PSYCHIATRIC DIAGNOSES ARE STILL UNRELIABLE

by Ahmed Aboraya, MD, DrPh

In 1980, the American Psychiatric Association (APA) Task Force, led by Robert Spitzer, developed and published the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III).[1] The DSM-III publication represented a benchmark in the history of psychiatric nomenclature because it included the long-awaited, detailed, explicit, and specific criteria of many psychiatric disorders.[2] Since 1980, the DSM-III and its subsequent editions have been used by psychiatrists and mental health professionals worldwide. The DSM-III was designed to serve both research and clinical purposes, and it did. Researchers use the DSM criteria to prove or elaborate on a particular hypothesis while mental health professionals use the DSM criteria when diagnosing patients in clinical practice. Even insurance companies require DSM diagnoses for reimbursement.[3] The DSM-III was also intended to improve the reliability of psychiatric diagnoses, an everlasting problem in psychiatry.[4–13]

Today, 26 years later, did the DSM system succeed in improving the reliability of psychiatric diagnoses? Two answers exist. The DSM did improve the reliability of psychiatric diagnoses at the research level. If a researcher or a clinician can afford to spend 2 to 3 hours per patient using the DSM criteria and a structured interview or a rating scale, the reliability would improve.[13] For psychiatrists and clinicians, who live in a world without hours to spare, the reliability of psychiatric diagnoses is still poor.[2,3] Even Spitzer and Frances, the directors of DSM-III and DSM-IV Task Force, admit that the desired reliability among the practicing clinicians has not been obtained.[3] To illustrate the problem of unreliablity, I reviewed the charts of a 64-year-old African American man who had more than 38 psychiatric admissions over a span of 43 years. Upon discharge the patient had the following diagnoses: schizophrenia, catatonia; schizophrenia, paranoid; schizophrenia, hebephrenic; schizophrenia, undifferentiated; schizoaffective disorder; bipolar type; and bipolar disorder with psychosis. Psychiatrists and clinicians attest that patients with multiple diagnoses are not uncommon.

The unreliability of psychiatric diagnoses is a complex topic and is more thoroughly explained elsewhere.[8–10,14–16] However, all diagnoses are affected by the following, which may account for unreliability:

1. Psychiatric nomenclature and classification
2. Patients' factors (anxiety, memory problems, defense

mechanisms)
3. Clinical presentations of psychiatric disorders (typical and atypical presentations)
4. Change of psychiatric symptoms over time for the same patient
5. Reliance on observing the patients' behaviors when they are unable to express their emotions
6. Reliance on proxy information in some cases
7. Clinician's style of interviewing
8. Clinician's experience
9. Clinician's bias toward certain diagnoses
10. Open-ended interview style and lack of methods to structure the clinician's interview
11. Clinician's training and school of thought
12. Constraints of time imposed on clinicians by institutions and financial incentives
13. Lack of agreement on definitions of psychiatric symptoms
14. Intentional change of diagnosis by clinicians for financial reasons (either to provide the patient with more services or to have insurance companies reimburse for services).

In summary, given the importance of having reliable diagnosis in modern psychiatry, more research and data are needed to explore the scope and causes of diagnostic unreliability in the clinical setting.

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Third Edition.* Washington, DC: American Psychiatric Press, Inc., 1980.
2. Aboraya A, Rankin E, France C, et al. The reliability of psychiatric diagnosis revisited: The clinician's guide to improve the reliability of psychiatric diagnosis. *Psychiatry 2006* 2006;3(1):41–50.
3. Spiegel A. The dictionary of disorder: How one man revolutionized psychiatry. *The New Yorker* 2005;56–63.
4. Andreasen NC, Grove WM, Shapiro RW, et al. Reliability of lifetime diagnosis. A multicenter collaborative perspective. *Arch Gen Psychiatry* 1981;38(4):400–5.
5. Beck AT, Ward CH, Mendelson M, et al. Reliability of psychiatric diagnosis. 2. A study of consistency of clinical judgments and ratings. *Am J Psychiatry* 1962;119:351–7.
6. Corty E, Lehman AF, Myers CP. Influence of psychoactive substance use on the reliability of psychiatric diagnosis. *J Consult Clin Psychol* 1993;61(1):165–70.
7. Goodman AB, Rahav M, Popper M, et al. The reliability of psychiatric diagnosis in Israel's Psychiatric Case Register. *Acta Psychiatr Scand* 1984;69(5):391–7.
8. Grove WM, Andreasen NC, McDonald-Scott P, et al Reliability studies of psychiatric diagnosis. Theory and practice. *Arch Gen Psychiatry* 1981;38(4):408–13.
9. Helzer JE, Clayton PJ, Pambakian R, et al. Reliability of psychiatric diagnosis. II. The test/retest reliability of diagnostic classification. *Arch Gen Psychiatry* 1977;34(2):136–41.
10. Helzer JE, Robins LN, Taibleson M, et al. Reliability of psychiatric diagnosis. I. A methodological review. *Arch Gen Psychiatry* 1977;34(2):129–33.
11. Kreitman N, Sainsbury P, Morrissey J, et al. The reliability of psychiatric assessment: An analysis. *J Ment Sci* 1961;107:887–908.
12. Mehlman B. The reliability of psychiatric diagnoses. *J Abnorm Psychol* 1952;47(2 Suppl.):577–8.
13. Skre I, Onstad S, Torgersen S, Kringlen E. High interrater reliability for the Structured Clinical Interview for DSM-III-R Axis I (SCID-I) *Acta Psychiatr Scand* 1991;84(2):167–73.
14. Carmines EG, Zeller RA. *Reliability and Validity Assessment.* London: SAGE, 1979.
15. Koran LM. The reliability of clinical methods, data and judgments (second of two parts). *N Engl J Med* 1975;293(14):695–701.
16. Koran LM. The reliability of clinical methods, data and judgments (first of two parts). *N Engl J Med* 1975;293(13):642–6.

If a researcher or a clinician can afford to spend 2 to 3 hours per patient using the DSM criteria and a structured interview or a rating scale, the reliability would improve. For psychiatrists and clinicians, who live in a world without hours to spare, the reliability of psychiatric diagnoses is still poor.

Dr. Aboraya is Clinical Associate Professor of Psychiatry at West Virginia School of Osteopathic Medicine, Director of the Young Adults Program at William R. Sharpe Jr. Hospital, and Director of the World Health Oganization (WHO) SCAN Training and Reference Center at West Virginia University, Morgantown, West Virginia.

Address correspondence to: Ahmed Aboraya, MD, 230 Canfield St., Star City, WV 26505; E-mail: aborayaahmed@wvdhhr.org

Given the crucial importance of the clinical utility of psychiatric diagnoses, studies should be conducted at the onset of the DSM-V revision process with the goal of shedding light on the clinical diagnostic process as currently practiced by mental health professionals.

COUNTERPOINT—THERE ISN'T ENOUGH EVIDENCE AVAILABLE TO SPECULATE ON THE RELIABILITY OF DIAGNOSES IN CLINICAL SETTINGS

by Michael B. First, MD

One of the most important goals of a psychiatric diagnostic system, such as the DSM, is to facilitate communication between clinicians. Diagnostic reliability is an important element in reaching this goal—clinical communication is undermined if two clinicians mean different things when they use the term *schizophrenia* to describe a particular patient's symptom presentation. It was for this reason (i.e., improving diagnostic reliability) that DSM-III introduced operationalized diagnostic criteria for every disorder in the classification.

While acknowledging that the reliability of psychiatric diagnoses has been improved at the research level, Dr. Aboraya in his column claims that "for psychiatrists and clinicians who live in a world without hours to spare, the reliability of psychiatric disorder is still poor," and he goes on to present 14 possible sources of such unreliability. On what data does he base this claim? He provides only two sources: 1) a *New Yorker* article, which quoted Robert Spitzer and Allen Frances (the respective chairs of the DSM-III and DSM-IV Task Force) as having doubts about reliability among clinicians,[1] and 2) his own paper that appeared in *Psychiatry 2006*.[2] A review of the references cited in his paper indicate that, with the exception of one 1989 Japanese study in which 20 psychiatrists made ratings on 28 case vignettes,[3] all of the studies cited as showing poor reliability were done prior to DSM-III, i.e., before the use of diagnostic criteria.

I am not claiming the contrary (i.e., that there is solid evidence that diagnostic reliability among clinicians since the advent of DSM-III is good). The fact is there is very little evidence available about the diagnostic reliability of the DSM system in clinical settings. The most comprehensive study of DSM reliability in clinical settings was the DSM-III field trials.[4] These field trials demonstrated good diagnostic reliability for most major classes of disorders, although these results have been called into question by critics of the DSM-III.[5] Since DSM-III, the field trials conducted under the auspices of DSM-III-R and DSM-IV have focused exclusively on testing proposed changes to specific criteria sets,[6,7] presumably reflecting the notion that DSM-III had "solved" the diagnostic reliability problem. Those reliability studies done since DSM-III have achieved good diagnostic reliability almost exclusively using structured interviews or checklists in which diagnostic raters methodically consider the presence or absence of each diagnostic criterion.[8] As Dr. Aboraya notes in his column, it is unlikely that clinicians "spend 2 to 3 hours per patient using the DSM criteria and a structured interview or rating scale" when making psychiatric diagnoses in clinical settings. Studies comparing clinical diagnoses with psychiatric diagnoses made using structured interviews have consistently shown poor agreement between these two methods,[9–12] suggesting that clinicians make diagnoses using different methods.

So how do clinicians make DSM-IV diagnoses in clinical settings? Remarkably, there have been virtually no studies published that have explored what clinicians actually do when they make psychiatric diagnoses.[13] Given this lack of fundamental baseline knowledge about clinician behavior, it is hard to speculate on what can and should be done to improve diagnostic reliability in clinical settings. Given the crucial importance of the clinical utility of psychiatric diagnoses,[13,14] studies should be conducted at the outset of the DSM-V revision process with the goal of shedding light on the clinical diagnostic process as currently practiced by mental health professionals. For example, in one such study, clinicians in different settings with various levels of experience fill out questionnaires at the conclusion of each diagnostic evaluation asking them to document precisely how they arrived at the diagnosis. Alternatively (or in addition to) *post-hoc* "debriefing sessions" using a focus group format could be conducted in which the

clinician's thinking processes are carefully examined. Once armed with this baseline information, we can then proceed with making informed decisions about how best to improve diagnostic reliability in clinical settings.

REFERENCES

1. Spiegel A. The dictionary of disorder: How one man revolutionized psychiatry. *The New Yorker* 2005 January 3, 2005:56–63.
2. Aboraya A, Rankin E, France C, et al. The reliability of psychiatric diagnosis revisited: The clinician's guide to improve the reliaiblity of psychiatric diagnosis. *Psychiatry 2006* 2006;3(1):41–50.
3. Kitamura T, Shima S, Sakio E, Kato M. Psychiatric diagnosis in Japan. 2. Reliability of conventional diagnosis and discrepancies with Research Diagnostic Criteria diagnosis. *Psychopathology* 1989;22(5):250–9.
4. Spitzer R, Forman J, Nee J. DSM-III field trials: I. Initial interrator diagnostic reliability. *Am J Psychiatry* 1979;136:815–7.
5. Kirk S, Kutchins H. *The Selling of DSM: The Rhetoric of Science in Psychiatry.* Hawthorne, NY: Aldine de Gruyter; 1992.
6. Spitzer RL, Siegel B. The DSM-III-R field trial of pervasive developmental disorders. *J Am Acad Child Adolesc Psychiatry* 1990;29(6):855–62.
7. Yutzy SH, Cloninger CR, Guze SB, et al. DSM-IV field trial: Testing a new proposal for somatization disorder. *Am J Psychiatry* 1995;152(1):97–101.
8. Skre I, Onstad S, Torgensen S, Kringlen E. High interrater reliability for the Structured Clinical Interview for DSM-III-R Axis I (SCID-I). *Acta Psychatr Scand* 1991;84(2):167–73.
9. Fennig S, Craig T, Tanenberg-Karant M, Bromet E. Comparison of facility and research diagnoses in first-admission psychotic patients. *Am J Psychiatry* 1994;151(10):1423–9.
10. Kashner T, Rush A, Suris A, et al. Impact of structured clinical interviews on physicians' practices in community mental health settings. *Psychiatr Serv* 2003;54(5):712–8.
11. Ramirez Basco M, Bostic JQ, Davies D, et al. Methods to improve diagnostic accuracy in a community mental health setting. *Am J Psychiatry* 2000;157(10):1599–605.
12. Shear M, Greeno C, Kang J, et al. Diagnosis of nonpsychotic patients in community clinics. *Am J Psychiatry* 2000;157(4):581–7.
13. First M, Pincus H, Levine J, et al. Clinical utility as a criterion for revising psychiatric diagnoses. *Am J Psychiatry* 2004;161(6):946–54.
14. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. *Am J Psychiatry* 2003;160:4–12.

Dr. First is Professor of Clinical Psychiatry, Columbia University, Research Psychiatrist, New York State Psychiatric Institute, New York, New York, and is a consultant to the American Psychiatric Association, publisher of the DSM-IV.

Address correspondence to: Michael First, MD, Professor of Clinical Psychiatry, Columbia University, Research Psychiatrist, New York State Psychiatric Institute 1051 Riverside Drive – Unit 60, New York, NY 10032; E-mail: mbf2@columbia.edu

INVITATION TO READERS—

In your opinion, are psychiatric diagnoses unreliable? If so, what are the important reasons for diagnostic unreliability among psychiatrists and clinicians?

The editors of *Psychiatry 2007* would like to know how our readers would answer this question. Your comments may be published in an upcoming issue of *Psychiatry 2007*.

Instructions for submitting your comments:

Please e-mail your comments to Elizabeth Klumpp, Executive Editor, eklumpp@matrixmedcom.com. Include "Response to Psychiatric Diagnoses" in the subject line of your e-mail. Limit your comments to 50 words. Please include your name, credentials (e.g., MD, PhD, etc.), title, affiliation, city, and state. There are no guarantees a comment will be published. Comments on this topic that are not intended for publication by the authors should state this in the body of the e-mail. Published comments may be edited for consistency of style and grammar and may be shortened. Include the following statement with the e-mail: "The undersigned author transfers all copyright ownership of the enclosed/attached letter to Matrix Medical Communications in the event the work is published. The undersigned author warrants that the letter is original, is not under consideration by another journal, and has not been previously published. I sign for and accept responsibility for releasing this material on behalf of any and all co-authors."